DOCUMENT RESUME

ED 040 113 SO 000 094

AUTHOR Bellack, Arno A.
TITLE Methods for Observing Classroom Behavior of Teachers and Students.
INSTITUTION Columbia Univ., New York, N.Y. Teachers College.
PUB DATE 12 Nov 68
NOTE 38p.; Presented at a conference sponsored by Padagogisches Zentrum, Berlin, November 12-15, 1968

EDRS PRICE EDRS Price MF-$0.25 HC-$2.00
DESCRIPTORS *Behavioral Science Research, *Classroom Observation Techniques, *Classroom Research, Educational Psychology, Educational Research, *Educational Sociology, Group Behavior, Group Dynamics, Interaction, Interaction Process Analysis, Research Design, Research Methodology, *Student Teacher Relationship, Teacher Evaluation, Test Selection, Test Validity

ABSTRACT

During the 60's, systems for the observation of classroom behavior have become an important area of educational research. Many new systems have been devised both for the evaluation of teaching techniques and for basic research into teaching processes. In the development of such systems the investigator must determine what behaviors are to be observed. This in turn requires such considerations as dimensions of classroom behavior, types of observation systems, and units of behavior. The second task of the investigator is to decide how the observations are to be carried out, e. g., recording and coding of behavior, reliability of coding, and statistical analysis. Three illustrations of observation systems are: Flander's system of classroom interaction analysis (affective dimensions), Bellack's observation of classroom discourse system (cognitive dimensions), and Oliver and Shaver's observational system for analyzing styles of teaching (affective and cognitive dimensions). The range of applicability of many observation techniques is limited and many investigators feel the need to devise their own. This has its drawbacks for comparisons even though the current state of the art often makes it necessary. (DJB)

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED040113

METHODS FOR OBSERVING CLASSROOM BEHAVIOR OF TEACHERS AND STUDENTS

Arno A. Bellack

Teachers College, Columbia University
New York, New York

This paper was written for discussion at the conference on "Methods of Determining Criteria for the Evaluation of Comprehensive Schools" sponsored by Pädagogisches Zentrum, Berlin, November 12-15, 1968. (To be published in report of the conference.)

SO 000 094

METHODS FOR OBSERVING CLASSROOM BEHAVIOR OF TEACHERS AND STUDENTS

Arno A. Bellack
Teachers College, Columbia University

A significant development in educational research during the past decade has been the burgeoning interest in the study of classroom teaching. This interest cannot accurately be called new; it represents rather a renewal of a long-standing concern of educational researchers. Among the features of contemporary studies of teaching that distinguish them from work of preceding periods is emphasis on systematic observation of the behavior of students and teachers in regular classroom settings. This paper presents a discussion of recent developments in methods for observing classroom behavior. The discussion is in three parts: first, an analysis of the task of constructing a system for observing classroom activities; second, brief descriptions of illustrative observation systems; and third, exploration of certain issues and problems in the development and use of systems for analyzing teacher and student behavior.

Development of Systems for Observing Classroom Behavior

In constructing a system for observing classroom behavior, the investigator must answer two basic questions: What dimensions of classroom behavior are to be observed? and How are the observations to be carried out? These two questions are, of course, intimately related; for <u>what</u> is observed and <u>how</u> one makes the observation are determined jointly. The

process of classroom observation is guided both by the researcher's conceptions of the behavior he is investigating and by the techniques and technology for gathering data available to him.

What Classroom Behaviors Are To Be Observed?

Since the researcher cannot possibly observe everything in the classroom that is there to be seen, he must decide what aspects of teacher and student behavior he is interested in studying. "Any observation is made," Kaplan writes; "it is the product of an active choice, not of passive exposure. . . . No one interpretation is necessitated by what is observed; there are always many ways of mapping behavior into data" (1964, p. 133). The way in which the researcher goes about mapping classroom behavior into data is dictated by the purposes of his study and the conceptual framework that informs and limits his observations. For example, Flanders' theoretical interest in "classroom climate" led him to develop a set of categories for describing contrasting modes of influence communicated in the teacher's verbal behavior (1965). In contrast, B. O. Smith and Meux's interest in the cognitive dimensions of classroom discourse led them to construct an observational schedule for categorizing logical operations in teaching (1962). And Kounin, because of his interest in classroom management and "discipline," devised an instrument for classifying dimensions of teacher style related to control of student behavior (1967).

To guide the observation of the dimensions of classroom behavior relevant to his purposes and theoretical orientation, the researcher is faced with the task of developing a reliable and valid instrument in the form of a set of rating scales or an observational system which specifies in operational terms the behaviors to be rated or categorized.

Rating Scales. The observer's task when using rating scales is to place the behavior being rated at some point along a continuum or in one of an ordered series of categories. (See Remmers, 1963, for a comprehensive discussion of rating methods in research on teaching.) An example of an experimentally developed numerical rating scale is that devised by Ryan (1960) in his research on teacher characteristics. In this research, the observer, after observing an entire class session, made assessments on each of 26 dimensions of the behavior that had occurred during that session and then recorded his assessments in the form of ratings on seven-point scales. The 22 dimensions of teacher behavior included partial-fair, harsh-kindly, aloof-responsive, autocratic-democratic and uncertain-confident; the four dimensions of student behavior included apathetic-alert, dependent-initiating, and obstructive-responsible. Each dimension was defined in a glossary which specified the relevant aspects of behavior upon which the rating was to be based.

Post-session rating procedures such as those developed by Ryan have serious limitations when used in the study of classroom activities. The assessments by raters provide only general impressions and recollections of what actually happened in the classroom, not accurate records of teacher and student behavior. For this reason, and because of certain deficiencies of rating scales as measuring instruments, few contemporary studies of teaching that involve direct classroom observation use them, although they have been widely used in the past.

Observation Systems. Classroom observation systems furnish the observer with a set of categories to which behaviors are assigned. Behavioral cues or operational definitions of the categories and coding instructions are frequently provided to help the observer decide in which

categories to classify observed behaviors. During recent years a large number of observation systems has been devised by researchers. In their anthology of observation instruments, Mirrors for Behavior, Simon and Boyer (1967) catalogued 26 systems, but these include no more than half of all recently developed systems.

There are many differences among existing observation systems: differences in (1) dimensions of classroom behavior to be classified, (2) type of observational schedule devised, (3) observer's frame of reference for coding, (4) unit of behavior to be used in coding, and (5) range of applicability. These differences highlight key problems in the construction of observation instruments.

(1) Dimensions of Classroom Behavior. As previously noted, a given system of analysis of necessity concentrates attention on certain facets of classroom behavior and omits others. Existing observation schedules are concerned with a wide range of behaviors. In a review of research on classroom behavior, Biddle (1967) found that recently devised instruments have dealt with (a) teacher performance in terms of actions, manners, and characteristic roles; (b) audience and target performance of teachers and students; (c) teacher-pupil interaction; (d) externally imposed structures, such as subject matter and administrative regulations; and (e) internal structures, such as communication structure, activity structure, characteristic roles, and social functions.

In contrast, Simon and Boyer (1967) have categorized observation schedules into three families of related systems: (a) affective systems dealing with the emotional climate of the classroom and how it is conditioned by teacher reactions to pupils' feelings, ideas, or actions; (b) cognitive systems concerned with thinking processes and verbal patterns used to deal with them; and (c) composite or multi-dimensional systems

dealing with both cognitive and affective dimensions of behavior. Flanders' system of interaction analysis (1965) and Hughes' system for classifying the functions of the teacher's classroom behavior (1959) are prominent examples of the first type. Bellack et al.'s scheme for describing the language of the classroom (1966) and B. O. Smith and Meux's system for analyzing logical operations of teaching (1962) are typical of the second group. Oliver and Shaver's observational schedule for describing contrasting styles of social studies teaching (1966) and Joyce's system for describing affective and cognitive aspects of the teacher's verbal communications (reported in Joyce and Harootunian, 1967) are representatives of the third type.

Existing classroom observation systems may also be classified on the basis of mode of communication -- that is, whether they focus attention on (a) verbal aspects, (b) nonverbal aspects, or (c) both verbal and nonverbal aspects of communication. By far the largest number of systems concentrate exclusively on verbal behavior, while a limited number deal with both verbal and nonverbal dimensions. Twenty of the systems described in Mirrors for Behavior (1967) are concerned with verbal communication, but only six give attention to both verbal and nonverbal behavior. To the present writer's knowledge, Galloway (1962), who studied nonverbal communications expressed in the teacher's gestures, quality of voice, and facial expressions, is the only researcher who has devised an instrument for describing nonverbal behavior per se.

Deciding what facets of teacher and student behavior to investigate is basically a substantive or conceptual problem. However, there are also important methodological considerations involved -- namely, making certain that the categories in a given system are mutually exclusive and exhaustive of a defined domain.

The criterion of exhaustiveness does not, of course, require that all behaviors during a given period of observation be classified. Rather, it requires that the domain or universe of behavior that the investigator is studying (e.g., social climate, logical operations, or communication processes) be clearly defined and characterized by an exhaustive set of categories. For example, the following five broad categories devised by Aschner and Gallagher (reported in Aschner, 1963) for classifying thought processes reflected in classroom verbal behavior are both mutually exclusive and exhaustive of the domain of cognitive functioning within their theoretical framework: routine, cognitive-memory, convergent thinking, evaluative thinking, and divergent thinking.

In the exploratory stages of research when the relevant theoretical concepts have not been fully developed, it is frequently difficult to devise categories that are exhaustive and mutually exclusive. However, as a line of investigation proceeds, the researcher must constantly keep these criteria in mind and give continuing attention not only to analysis of empirical data of observation but also to refinement of the concepts he uses to analyze these data.

(2) Types of Observation Systems. Two types of observation systems are widely used in contemporary classroom research, and they have been classified by Medley and Mitzel (1963) as "category systems" and "sign systems."

A category system limits observation to specified dimensions of classroom behavior, providing a set of categories into which each unit of observed behavior is classified. The resulting record shows, for each period of observation, the total number of units of behavior that occurred and the number classified in each category. The set of categories developed

by Hughes (1959) in her study of interaction in the elementary classroom is typical. The observer using this system classifies the teacher's verbal and nonverbal behavior in terms of six major "functions" it serves for students: controlling functions, facilitating functions, functions that develop content, functions that serve as personal response, functions of positive affectivity, and functions of negative affectivity.

In contrast, a sign system provides the observer with a listing of specific behaviors which may or may not occur during a period of observation. Observers are instructed to be on the lookout for these specific behaviors, and the record obtained shows which of these behaviors occurred during the period of observation. An example of this type of system is the Observation Schedule and Record (OScAR) developed b Medley and Mitzel (1958). An observer using this instrument records only those behaviors that fall into one of the 71 items listed. These items, grouped into several sections, include such signs as "teacher lectures," "teacher answers pupil's questions," "teacher illustrates at the board," "teacher uses sarcasm," "pupil talks to group," "pupil whispers," and "pupil reads or studies at his seat."

Whereas sign systems usually consist of a large number of items referring to concrete, specific behaviors and therefore requiring little inference on the part of the observer, category systems usually are made up of a smaller number of items at a higher level of abstraction which demand a higher degree of inference by theobserver. All category systems are not constructed at the same level of abstractness, but as a rule they tend to be found at higher levels of conceptualization than sign systems. The examples cited in the two preceding paragraphs illustrate the difference between high-inference categories and low-inference signs. It requires

very little inference on the observer's part to interpret a teacher's action as an instance of "answering pupil's question" or "illustrating on the board." In contrast, it demands a higher degree of inference to interpret a teacher's statement as serving a controlling function or a facilitating function. It should also be noted that high-inference measures concerned with observations of behavior as a basis for inferences about motives of individuals or about effects of behavior present the researcher with more difficult problems in estimating validity than do low-inference measures which are concerned with observations of behavior used descriptively and thus are often said to have "face validity."

Medley and Mitzel (1963) observe that category systems have been used more frequently in studies based on well-developed theories, while sign systems have been used when theory has not provided much guidance. Clearly, the theoretical orientation of the researcher and the stage of development of his theory dictate the level of conceptualization and therefore the degree of inference required.

(3) Observer's Frame of Reference. Researchers attempt to categorize classroom events from at least three perspectives: in terms of (a) the intent or motive of the actor, (b) the effects of the behavior on the recipient(s) of the action, or (c) the objective characteristics of the behavior. Examples of these three perspectives may readily be identified in recent studies of classroom processes.

Withall's system for analysis developed in connection with his studies of the social-emotional climate of the classroom (1949) requires the observer to interpret teachers' statements as to whether their intent is pupil-supportive, problem-structuring, or directive. Hughes (1959), on the other hand, in her investigation of the classroom behavior of elementary school teachers, classified the "functions" of the teacher's

verbal and nonverbal behavior in terms of their inferred effects on, or significance for, students. In contrast, Smith and Meux (1962) in their study of the logical aspects of classroom discourse describe the logical features of teachers' and students' statements, making no assumptions about motives of the speaker or effects on the audience.

The experiences of these and other researchers demonstrate that classroom behavior can be coded reliably from all three perspectives. Which of the three frames of reference is appropriate for a given study can be determined only by considering the objectives of that study. Biddle (1967) has identified several possible purposes that classroom research might be designed to serve and "conceptual postures" appropriate for these purposes. For example, he suggests that if one is primarily interested in the determinants of teacher behavior, then judgments of teacher intent are appropriate. If, in contrast, one is concerned with teacher competence, then judgments of the effects of teacher behavior on student learning would be more appropriate. If, on the other hand, one is interested in studying both individual and social determinants of behavior or in testing contrasting models of classroom interaction, then it would be well to emphasize objective characteristics of behavior.

(4) Unit of Behavior. A crucial task in devising an observation system is specifying the unit of behavior that is to be used as the basis for coding. The variety of ways in which researchers have defined the unit of behavior may be classified into two principal approaches: (a) designation of an arbitrary time unit; and (b) specification of an analytic unit, frequently one suggested by the category system itself.

In those systems in which an arbitrary time unit is employed, the observer is required to make a record of what behavior or behaviors

occurred during the specified time. For example, Flanders (1965) requires the observer to make a judgment every three seconds as to whether the teacher during her brief period is exerting direct or indirect influence on students. Similarly, Spaulding (reported in Simon and Boyer, 1967) specifies time periods ranging from three to ten seconds during which the observer makes a judgment regarding the teacher's approach to control in the classroom, using such categories as "setting performance goals" and "prescribing" certain kinds of activities. Medley and Mitzel (1958), on the other hand, expect the observer to keep a record of signs for five-minute periods.

The principal advantage of the arbitrary time unit lies in its automatic character which helps to regularize the process of observation. Furthermore, it can be used in analyzing both verbal and nonverbal behavior. However, the difficulty with the time unit is that when classroom events are tallied at arbitrary time intervals, the resulting data do not reflect the naturally occurring pattern or stream of behavior as it develops over time.

Instead of specifying an arbitrary time unit, many researchers devise analytic units of various types. Analytic units represent discrete elements of verbal and/or nonverbal behavior that are used in coding to divide classroom events into component parts. Researchers have defined analytic units in a variety of ways: in terms of (a) classroom activity, (b) verbal transaction between two or more speakers, (c) communication or message of an individual speaker, (d) "item of thought" expressed by a speaker. These types of analytic units may be illustrated by the following examples:

(a) In his study of classroom behavior and underachievement, Perkins (1964, 1965) specified as the basic units of analysis six types of class activity: large-group discussion, class

recitation, individual work or project that is not a common assignment, seatwork on common assignment, small-group or committee work, and oral reports. Aspects of pupil and teacher behavior were coded in the context of these types of activity.

(b) In their study of the logical dimensions of classroom discourse, B. O. Smith and Meux (1962) designated the basic behavioral unit as "the episode" which was defined as a verbal exchange between two or more speakers consisting of three phases -- an initial or opening phase, a continuing phase, and a terminal phase.

(c) Jackson (1965) defined three types of "verbal messages" as the basic units for describing communication in the elementary classroom: instructional messages referring to content objectives, group management messages having to do with procedures and rules, and control messages concerned with maintaining discipline and keeping order.

(d) In her study of teaching strategies for teaching cognitive skills to elementary school children, Taba, et al. (1964, 1966) specified the basic unit of analysis as the "thought unit" which was defined as "a remark or series of remarks which expressed a more or less complete idea, served a specific function, and could be classified by a level of thought" (1966, p. 134).

(5) Range of Applicability. Observation systems differ in the degree to which they are applicable to research settings and populations other than those for which they were originally constructed.

Some systems were designed with the purpose of making them useful in many different types of classrooms at various grade levels, in various subjects and involving various types of students. For example, Flanders' system of interaction analysis (1965) is applicable in both elementary and secondary classrooms in which teachers interact verbally with students. Similarly, Biddle and Adams' scheme for analyzing the structural and functional features of the classroom communication system (1967) was developed with the deliberate objective of making it applicable in elementary and secondary classrooms in which a variety of subjects are taught.

In contrast, some observer systems are appropriate for analyzing classroom processes at certain grade levels. Hughes' system for describing the functions of teachers (1959) is applicable to classrooms at the elementary level, while B. O. Smith and Meux's system for analyzing logical operations in teaching (1962) is relevant only to high school classes dealing with academic subjects.

Still other systems are limited to certain subjects or to certain types of students. Wright's system (1961) is designed for studying mathematics classes, while Oliver and Shaver's scheme (1967) is applicable primarily to social studies classrooms. Some researchers are concerned only with certain types of students: Kounin (1967) with emotionally disturbed pupils, Perkins (1964, 1965) with underachieving students, and L. Smith and Geoffrey (1968) with students representing lower socio-economic classes.

Whether researchers should attempt to develop comprehensive multi-dimensional systems applicable to a wide range of classroom settings or concentrate on building instruments designed for a more limited range of settings is a debatable issue to which we shall return later in this paper.

How Are Classroom Observations To Be Carried Out?

Recording and Coding Behavior. A variety of different techniques for recording and coding behavioral data has been used in connection with the observation systems discussed above. Some investigators (e.g., Flanders, 1965; Jackson, 1965; and Medley and Mitzel, 1959) require the observer to code behavior as it actually occurs in the classroom; that is, the "observer-coder" transforms the observed events into symbols that may be counted and tabulated. The principal advantage in coding behavior "on the spot" is that the coder has direct access to visual cues such as facial expressions and gestures of teachers and students as well as situational factors in the classroom that may be relevant in interpreting accurately the behavior under observation. The most serious disadvantage is that it is exceedingly difficult to code reliably complex behavioral processes at the rapid pace at which they occur in the classroom. Many researchers, therefore, follow the procedure of first recording classroom behavior and then coding the behavior on the basis of the recordings.

Various types of behavioral recordings have been used in studies of the classroom, including specimen recordings, audio recordings, and audio-visual recordings. In her study of elementary teaching, Hughes (1959) collected data in the form of specimen records which were essentially sequential narrative accounts of the teachers' verbal and nonverbal classroom behavior. Two trained observers simultaneously recorded events in the classroom, and the final specimen record of a given period of observation included only descriptions that both observers agreed upon. No attempt was made to categorize behavior; observers merely recorded in shorthand what teachers said and did. Coders later classified the recorded data in terms of seven major categories of teaching functions.

Hughes contends that gathering classroom data in the form of specimen records has several advantages including the following: This procedure enables the researchers to "hold teaching still" in a permanent form so that it may be studied and its distinctive quality and characteristics identified; the continuity of the teacher's behavior is retained; and the records are neutral in that no judgment is made regarding what is taking place. Serious questions may be raised regarding the assumed "neutrality" of specimen records, for these records are in fact "second-hand" accounts of classroom behavior by the observer whose perceptual biases are inevitably reflected in his record of what transpired.

This shortcoming of specimen recording can largely be overcome through the use of sound recordings and audio-visual recordings which provide permanent, objective records of behavior. In several studies investigating classroom verbal behavior (e.g., Bellack, et al., 1966; Taba, et al., 1964, 1966; and B. O. Smith and Meux, 1962) data have been collected through the use of electronic recording equipment. Frequently, typewritten records are made from the sound recordings, and the coder then has access both to tape recordings and to typescripts when categorizing the behavior. In view of the difficulty of observing and analyzing complex verbal behavior as it actually occurs in the classroom, there are obvious advantages in securing a permanent sound recording and typescript of the discourse which can be repeatedly observed and analyzed from a variety of different perspectives. The principal shortcoming of sound recordings is that they provide no information regarding nonverbal behavior which might be significant in itself, or might serve as cues for the accurate interpretation of verbal behavior.

This difficulty may be overcome through the use of audio-visual recordings which, since they record both verbal and nonverbal behavior, provide a comprehensive record of classroom events. Kounin (1967) and Biddle and Adams (1967) have recently completed studies in which data were gathered in regular classroom settings using portable (and expensive!) videotape equipment. The principal problems are the costs involved and technical difficulties in using the complicated equipment. The reader interested in the use of videotapes in classroom research is referred to the report of Biddle and Adams' research (1967, Chapter IX).

Relation of Observer and Observed. A frequently expressed objection to observation studies is that the presence of the observer or recording equipment is so distracting that the observed behavior cannot be regarded as "typical" behavior. Fortunately, this does not seem to be a valid criticism, if one can judge from the opinions of experienced researchers. In their extensive review of systematic observation techniques in social psychological research, Heyns and Lipett report that experienced users of observers share the common feeling that "observers have very little effect, if any. This belief is shared by experimenters who have worked in a wide variety of situations and with many kinds of subjects" (1954, p. 399).

Most educational researchers seem to share this opinion. For example, Biddle and Adams (1967), whose complicated audio-visual equipment included two cameras and two microphones placed in the classrooms they were studying, comment as follows about the effect of this equipment on teachers and pupils:

> All participating teachers were interviewed informally about this [i.e., effect of recording equipment]. Some reported some feeling of tension at the beginning of the first recorded session but testified that it disappeared as they became involved in the lesson. The researchers noted that several teachers appeared to "dress up" for re-

> cording days, although no such effect was noted for pupils who -- it will be recalled -- did not know on which day recordings were to be made. Occasionally, pupils would give evidence of being aware of the recording cameras -- gazing at them speculatively or, at interval time, 'performing' in front of them. However, the overall impression gained was that the cameras overtly distracted the pupils to hardly any extent at all (p. 217).

Similar opinions regarding negligible effects of observers and recording equipment on classroom behavior have been expressed by other researchers, including B. O. Smith and Meux (1962), Hughes (1959), Flanders (1965), and Bellack, et al. (1966).

Researchers have found certain measures helpful in minimizing the effects of observers and recording equipment on classroom behavior. For example, B. O. Smith and Meux (1962) indicate that instructions to their cooperating teachers assured them that the research did not involve an evaluation of their teaching performance, that complete anonymity would be maintained in reporting results of the research, and that only research staff members would have access to the tapes and typescripts. Furthermore, it has become common practice to install recording equipment in the classroom a few days in advance of the experimental sessions to enable students and teachers to become accustomed to it.

Without minimizing the problems posed by the effect of theobserver and recording equipment, it would seem that these problems can to a large extent be met if precautionary measures are taken. While it cannot be assumed that teachers and students will be unaffected by the presence of an observer or recording equipment even when such measures are taken, it is well to remind ourselves of the simple and obvious point that "to know how teachers and pupils behave while they are under observation seems better than to know nothing at all about how teachers and pupils behave" (Medley and Mitzel, 1963, p. 248).

Reliability of Coding. Reliability of a classroom observation instrument may be defined as "a measure of the extent to which a measurement remains constant as it is repeated under conditions taken to be constant" (Kaplan, p. 200). Among the relevant conditions in measuring classroom behavior, the observers making the measurement and the stability of the behavior being measured are of particular importance. Therefore, reliability of instruments for observing classroom events may be estimated in terms of degree of agreement between independent observers (coefficient of agreement) and in terms of stability of the dimensions of behavior under observation (stability coefficient). Although both types of reliability are obviously important, relatively little effort is made in contemporary research on classroom behavior to estimate stability coefficients; major attention is focused on estimating the degree of agreement between observers. The more common statistical indices used in making these estimates are percent of agreement between coders and the correlation coefficient, but analysis of variance also is used by some researchers.

Researchers have found certain procedures helpful in insuring adequate reliability in the use of observation schedules. Of central importance is painstaking development of the observation schedule itself, giving particular attention both to precise definition of the categories and the unit of analysis, and to formulation of coding rules for the guidance of coders. Probably the most serious difficulties in coding reliability result from lack of clarity in definition of the unit of behavior that is to be coded; without agreement on the basic behavioral unit to be analyzed, it is obvious that a high degree of inter-observer agreement is impossible. Careful training of observers is, of course, a principal means for insuring a high degree of reliability; the experience of researchers bears testimony

to the importance of intensive training programs for observers and coders, and systematic procedures for checking the reliability of their observations and coding.

Statistical Analysis of Data. Detailed discussion of methods of statistical analysis appropriate for classroom observational data is beyond the scope of this paper. However, it is worth noting that recently developed statistical theories and techniques make it possible to investigate certain dimensions of classroom processes that previously have been difficult, if not impossible, to study statistically. For example, Biddle and Adams (1967) and Bellack, et al. (1966) have statistically described the temporal patterning of classroom events through a Markov chain. In the latter study (Bellack, et al., Chapter VII), the researchers sought to determine whether certain cyclical patterns of pedagogical moves tend to influence subsequent patterning of pedagogical moves. Statistically this was described through a Markov chain, which makes it possible to determine the transition probabilities of moving from one state to another (Kemeny and Snell, 1962). Taking types of teaching cycles (patterns of pedagogical moves) as states, the probabilities of moving from one type of cycle to another were investigated as a way of determing whether one pattern of pedagogical moves tends to influence immediately subsequent patterning.

Statistical treatment of data in observational studies has been greatly facilitated through the use of high-speed electronic computers. A significant recent development in the use of computers has been reported by Gage (in press) in the work of Allen and Snow at the Stanford Center for Research and Development in Teaching. Allen and Snow have begun the development of a generalized taxonomy of classroom behaviors, using computers for storage and retrieval of items referring to teacher

and learner behaviors. Gage describes the purposes and procedures of this project:

> More than 1,000 items have been stored; they can be retrieved and printed by the computer in a form ready for use by observers, raters, and content analysts. The items to be retrieved can be specified according to many different dimensions. Data generated by observers or content analysts concerning the frequency, intensity, or correlates of behaviors denoted by these items can be stored in the computer along with the items. Thus, experience relevant to the reliability and validity of the items with different kinds of teachers, observers, subject matters, criteria of effectiveness, and the like, can become cumulative. In short, the ideal of a universal taxonomy of classroom behaviors useful for many alternative purposes may become attainable through a computer-based system like that being developed at Stanford.

Illustrative Observation Systems

This section presents brief descriptions of three systems of analysis that are illustrative of three major types of existing observation systems: (1) Flanders' system of interaction analysis, an affective system dealing with the social climate of the classroom (1965); (2) Bellack, et al.'s system for analyzing the language of the classroom concerned primarily with cognitive dimensions of teaching (1966); and (3) Oliver and Shaver's system for describing teaching styles in social studies classrooms, a multi-dimensional system which focuses attention on both affective and cognitive aspects of classroom activities (1966). Since all three category systems may be used in studying secondary classrooms, they were thought to be appropriate for discussion at this conference.

A System For Analyzing Affective Dimensions of Teaching

Conceptual Framework. Flanders' system of classroom interaction analysis (1965) focuses primary attention on affective and interpersonal components of classroom processes. Flanders contends that the way in

which the teacher, through his verbal communications, seeks to influence student behavior is the most important feature of teacher-student relationships in the classroom. His research efforts are directed toward analyzing two contrasting methods of influence used by the teacher: (a) direct influence which "consists of those verbal statements of the teacher that restrict freedom of action, by focusing attention on a problem, interjecting teacher authority, or both;" and (b) indirect influence "consists of those verbal statements of the teacher that expand a student's freedom of action by encouraging his verbal participation and initiative" (p. 9).

Categories and Method of Analysis. Flanders' category system, designed for classroom situations in which teacher and students are actively engaged in verbal interaction, describes verbal acts of the teacher that exert direct and indirect influence. There are ten items in the observational scheme as indicated in the following chart.

FLANDERS' CATEGORIES FOR INTERACTION ANALYSIS

TEACHER TALK	Indirect Influence	1. Accepts Feeling
		2. Praises or Encourages
		3. Accepts or Uses Ideas of Student
		4. Asks Questions
	Direct Influence	5. Lecturing
		6. Giving Directions
		7. Criticizing or Justifying Authority
STUDENT TALK		8. Student Talk - Response
		9. Student Talk - Initiation
		10. Silende or Confusion

Seven of the categories are used for classifying teacher talk, two for student talk. The tenth category covers pauses, short periods of silence, or confusion. Of the seven categories assigned to teacher talk, items 1 through 4 represent indirect influence, and items 5 through 7, direct influence.

Data are gathered by an observer in the classroom. At the end of each three-second period, the observer decides which of the ten categories best describes the communication events of that time period. He writes this number down while simultaneously assessing communication in the next three-second period. He continues at a rate of about 20 observations per minute. His notations are a sequence of numbers written in a column, top to bottom, so that the original sequence of events is preserved. When there is a shift in classroom activities, the observer draws a double line and indicates the time. Identifying activity-periods in this manner is thus a second mode of categorization that is superimposed on the system for classifying verbal statements. Flanders identifies five types of activity: introducing new material; evaluating homework or tests; other class discussion; supervising seatwork; and routine seatwork.

After a period of observation, the observer converts the numbers he has listed in sequence to tallies in a 10 x 10 matrix. The numbers are tallied in the matrix one pair at a time. For example, if the observer has recorded the series 10, 6, 7, 5, 4, 8, 1, 4 and 8, the first pair is 10-6; the second pair, 6-7; the third pair, 7-5; etc. For the first pair, 10-6, the tally is placed in the row 10, column 6 cell; the second pair, 6-7, in the row 6, column 7 cell; etc. This type of analysis helps preserve, to some extent, the sequential nature of classroom events. Analysis of the matrix data provides various kinds of information about the inter-

action process in the classroom: how much the teacher talks, how much the students talk, the amount of time spent in silence or confusion, the proportion of indirect and direct influence exerted by the teacher (referred to as the I/D ratio), the sequences of types of teacher and student talk, and the like.

Reliability. Flanders and his co-workers have developed systematic procedures for training observers to insure adequate reliability of coding. Inter-observer reliability is estimated by using Scott's coefficient. In a recent study reported by Flanders (1965), Scott reliabilities were consistently above 0.85 for trained observers.

Range of Applicability. This system of interaction analysis has been widely used in a variety of descriptive and experimental investigations. Recently, it has also been used in pre-service and in-service training of teachers to enable them to obtain feedback about their own classroom behavior.

A System for Analyzing Cognitive Dimensions of Teaching

Conceptual Framework. In their study, The Language of the Classroom (1966), Bellack and his associates conceptualized classroom discourse as a kind of "language game," following Wittgenstein's view of language as rule-governed verbal behavior. The basic unit of discourse for describing the classroom game is "pedagogical move." Moves are classified in four major categories according to the pedagogical functions they serve in classroom discussion:

> Structuring. Structuring moves serve the pedagogical function of setting the context for subsequent behavior by either launching or halting-excluding interaction between students and teachers. For example, teachers frequently launch a class period with a structuring move in which they focus attention on the topic or problem to be discussed during that session.

Soliciting. Moves in this category are designed to elicit a verbal response, to encourage persons addressed to attend to something, or to elicit a physical response. All questions are solicitations, as are commands, imperatives, and requests.

Responding. These moves bear a reciprocal relationship to soliciting moves and occur only in relation to them. Their pedagogical function is to fulfill the expectation of soliciting moves; thus students' answers to teachers' questions are classified as responding moves.

Reacting. These moves are occasioned by a structuring, soliciting, responding, or prior reacting move, but are not directly elicited by them. Pedagogically, these moves serve to modify (by clarifying, synthesizing, or expanding) and/or to rate (positively or negatively) what has been said previously. Reacting moves differ from responding moves: while a responding move is always directly elicited by a solicitation, preceding moves serve only as the occasion for reactions. Rating by a teacher of a student's response, for example, is designated as a reacting move.

Moves, which describe the verbal behavior of both teachers and pupils, occur in classroom discourse in certain cyclical patterns or combinations which are designated "teaching cycles." For example, a typical teaching cycle consists of a teacher solicitation followed by a pupil response which in turn is followed by the teacher's reaction to the responding move. The sequential ordering of teaching cycles in classroom discourse is described through a Markov chain as described previously on page 18.

Analysis of the classroom game would be incomplete without description of the content of the messages communicated in the moves made by participants. Two basic types of content are identified: (a) substantive meanings which refer to the subject matter under discussion -- in this research, the topic of "international trade:' and (b) instructional meanings which refer to assignments, materials, and classroom procedures. The substantive and instructional meanings are observed and recorded along with their associated logical meanings which refer to cognitive processes

involved in dealing with substantive and instructional content such as defining, interpreting, fact stating, explaining, opining, and justifying. Thus, if a pupil answers a teacher's question by giving a definition of tariff, the pedagogical move of the pupil is coded as a responding move, the substantive meaning as tariff, and the logical process as defining. In addition, the pupil is designated as the speaker, and the length of his move in lines of typescript is also recorded.

Coding Procedures. Coding is done from the viewpoint of the observer, with pedagogical meaning inferred from the speaker's verbal behavior. Coders listen to tape recordings of class sessions and also follow transcribed protocols. Each pedagogical move is coded according to the categories of analysis summarized as follows:

(1) Speaker
(2) Type of Pedagogical Move
(3) Substantive Meaning
(4) Substantive-Logical Meaning
(5) Number of lines in (3) and (4)
(6) Instructional Meaning
(7) Instructional-Logical Meaning
(8) Number of lines in (6) and (7)

An example of a coded pedagogical move is:

T	/	STR	/	IMX	/	XPL	/	4	/	PRC	/	FAC	/	2
(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)

This is interpreted as follows: a teacher makes a structuring move in which he explains something about imports and exports for four lines of transcript and also states facts about class procedures for two lines of transcript.

The following is a brief summary of the categories of analysis, with the eight numbers corresponding to the numbers given above:

1. Speaker
   - Teacher (T)
   - Pupil (P)
   - Audio-visual device (A)

2. Type of Pedagogical Move
   - Structuring (STR)
   - Soliciting (SOL)
   - Responding (RES)
   - Reacting (REA)

3. Substantive Meaning (Subject matter which in the case of this research was "International Trade;" categories based on content analysis of pamphlet used in participating classes)

   - Trade (TRA)
   - Factors of Production and/or Specialization (FSP)
   - Imports and/or Exports (IMX)
   - Foreign Investment (FOR)
   - Barriers to Trade (BAR)
   - Promoting Free Trade (PFT)
   - Relevant to Trade (REL)
   - Not Trade (NTR)

4. Substantive-Logical Meanings (Cognitive processes involved in dealing with subject matter)

   Analytic Processes
   - Defining (DEF)
   - Interpreting (INT)

   Empirical Processes
   - Fact Stating (FAC)
   - Explaining (XPL)

   Evaluating Processes
   - Opining (OPN)
   - Justifying (JUS)

5. Number of lines of typescript in (3) and (4)

6. Instructional Meanings (Factors related to classroom management)

   Assignment (ASG)
   Material (MAT)
   Person (PER)
   Procedure (PRC)
   Statement (STA)
   Logical Process (LOG)
   Language Mechanics (LAM)
   Action-General (ACT)
   Action-Vocal (ACV)
   Action-Physical (ACP)
   Action-Cognitive (ACC)
   Action-Emotional (ACE)

7. Instructional-Logical Meanings (Cognitive processes associated with instructional meanings)

   Analytic, Empirical, Evaluative Processes
   (same as 4 above)

   Rating (Reference to metacommunication, usually an evaluative reaction)

   Positive (POS)
   Admitting (ADM)
   Repeating (RPT)
   Qualifying (QAL)
   Not Admitting (NAD)
   Negative (NEG)

   Extra-Logical Processes

   Performing (PRF)
   Directing (DIR)

8. Number of lines in (6) and (7) above

Reliability. To estimate reliability of the coding system, two teams of the research staff coded samples of transcript selected at random, and compared results. The percentage of agreement was calculated for each of the basic categories of the system of analysis in terms of number of moves and number of lines. The results indicated consistently high degree of reliability for all categories of analysis: agreement ranged from 84 to 96 percent.

Range of Applicability. With appropriate modifications in the categories labeled "substantive meanings," this system provides a potentially useful technique for studying classroom behavior in a variety of subjects at both the elementary and secondary levels. Doctoral dissertations recently completed at Teachers College have adapted this method of analysis to the study of mathematics teaching at the junior high school level and the teaching of reading at the elementary level.

A System for Analyzing Affective and Cognitive Dimensions of Teaching

Conceptual Framework. In connection with their study, Teaching Public Issues in the High School (1966), Oliver and Shaver devised an observational system for analyzing two styles of teaching: recitation teaching and socratic teaching. The theoretical framework within which these two types of discourse are defined involves three dimensions of classroom interaction: (a) cognitive dimensions, (b) affective or socioemotional dimensions, and (c) procedural dimensions.

With respect to the cognitive aspects, "recitation teaching tends to be descriptive; it is assumed that the truth of the situation is available and that one has only to present and clarify information or an analytic structure by which information can be organized" (p. 178). In contrast, "The socratic style is clearly dialectical. It assumes that the problem can be clarified only in an adversarial context, in which various points of view are presented and defended" (p. 178). With respect to affective aspects of interaction, it is suggested that both socratic and recitation teachers need to be supportive of students. However, "because the open controversy on the cognitive level 'spills' over into the affective domain,

the socratic discussion will tend to be highly charged with negative affect" (p. 179). Procedurally, since the teacher is the focus of the dialogue between himself and the students, no systematic differences between recitation and socratic teachers are expected in this dimension.

Categories of Analysis. Oliver and Shaver created a set of categories that have meaning in terms of recitation and socratic styles of teaching as described above. Their experimental instrument consists of three sets of categories, as indicated below.

Affective or Socioemotional Categories

1. Solidarity
2. Low Positive Affect
3. Tension Release
4. Tension
5. Low Negative Affect
6. Antagonism

Neutral (No affective message)

Cognitive Categories

7. Suggests Inconsistency
8. Descriptive
9. Evaluation
10. Repeats, Summarizes, Focuses
11. Clarification
12.. Analogy

Non-cognitive

Procedural Categories

13. Directs Task-oriented Behavior
14. Controls Deviant Behavior

The observer using this instrument must infer the cognitive or procedural significance in each statement or "single item of thought" expressed by the teacher. Each such scorable act is categorized in a cognitive or procedural category and in an affective category. The functions of the cognitive categories are described as follows:

> The primary function of the cognitive categories (7-12) is to answer questions about differences between the intellectual or logical content of discussions led by teachers using the two teaching styles. Central to the distinction between the two styles is the extent to which teachers deal with descriptive information in the controversial case as opposed to value judgments arising from the cases. Categories 8 and 9 are set up specifically to identify differences of this kind. Category 7 (Suggests Inconsistency) is meant to identify attempts by the teacher to arouse personal value conflicts on the part of the student by suggesting that he is making contradictory judgments in similar situations (the case and an analogy). Category 12 (Analogy) thus has obvious significance. Categories 10 and 11 were included to make the cognitive subsystem exhaustive (pp. 291-292).

Reliability. Observers were carefully trained in the system, using tape recordings of class discussions. An adaptation of Chi Square was used to estimate inter-observer agreement. All of the Chi-Square values were well below the .50 probability level criterion which the researchers adopted.

Range of Applicability. Although this system of analysis was devised for use in social studies classes in which public issues are the focus of discussion, the authors suggest that it might be adapted for the analysis of teaching in other subjects, such as science and literature, as well.

Problems and Issues

In this section the discussion turns to two of the many problems and issues that educators face in developing and using classroom observation systems.

Relationships Among Observer Systems

The first problem grows out of the current tendency among researchers to devise a new system of analysis for virtually each new study, apparently on the assumption that systems developed by others are inappropriate or inadequate. It may indeed be that a given research problem deals with facets of classroom behavior for which an observational schedule has not yet been developed, or that available systems which might be relevant are not valid or reliable. However, the net result of the proliferation of observation instruments has been that researchers are gathering very little data that are comparable from one study to another, making the task of comparing and contrasting the findings of these studies exceedingly difficult.

The problem of the proliferation of observation systems is undoubtedly the result of competing conceptual orientations of researchers. In this connection, it is well to keep in mind that "in the early stages of the development of any science different men confronting the same range of phenomena, but not usually all the same particular phenomena, describe and interpret them in different ways" (Kuhn, p. 17). Thus workers in the field of research on teaching -- a field that just now is establishing its foundations -- cannot be expected to pursue their research from a common theoretical orientation. At this stage of the game, conflicting and competing approaches to the description of classroom events are to expected and even encouraged.

However, this state of affairs does not preclude the possibility of investigations to determine basic factors or dimensions underlying the various observation systems so that differences and similarities among

them can be more clearly identified. A factor analytic approach to comparing observation systems along basic underlying dimensions for the purpose of identifying differences between them has been proposed by Gage (in press):

> The behaviors of a large sample of teachers would be measured on many variables of the kinds that have been specified by Flanders (1964), Smith, et al. (1967), Bellack, et al. (1966), Spaulding (1965), Medley and Mitzel (1959), and others. Then the intercorrelations of the scores on the variables would be subjected to factor analysis. The resulting factors would define the dimensions in relatively parsimonious, or "simple structure," terms.

Once the results of comparative and empirical studies of existing systems become available, the way is open to the construction of multi-faceted observation schedules that incorporate concepts from a variety of different systems. Meux (1967), for example, has proposed the development of a new multi-aspect system to include topic-objective units, initiating-reacting cycles within the topic units, content-substantive and instructional facets, group processes, and thinking processes. He argues

that if categories such as these from several different observation systems are combined into the multi-dimensional system proposed, teaching strategies might be constructed which are far more effective than those presently used by our best teachers. "This possibility of new combinations of present practices has been overlooked," he suggests, "by a number of writers who argue that observation studies can in principle never get us beyond present practice" (pp. 549-550).

Along different lines, but with similar purposes in mind, Biddle (1967) has suggested the possibility of an "ideal system" which would include three broad sets of concepts: (a) concepts for classroom activities as a whole, including structural and functional aspects of the social system represented by the classroom; (b) concepts for public discourse of interactive moves made by classroom participants, and (c) concepts for classroom language usage, including syntax and phonology.

Whether at this stage in the development of classroom research, researchers should give major attention to the construction of multi-dimensional systems like those suggested by Meux and Biddle, or whether primary effort should be directed to building observation systems of more limited scope is a debatable question. There seems to be little reason, however, to propose that researchers limit themselves to one approach or the other. Given our limited knowledge about the teaching process, one might reasonably expect both approaches to contribute to our understanding of classroom events.

Potential Uses of Observer Systems

Observation of teacher and pupil behavior has come to be an accepted method for collecting data in investigations of classroom teaching. Data

provided through observational techniques give promise of increasing our understanding of the complex life of the classroom. Increasingly widespread use of observational techniques in research on teaching has stimulated educators concerned with teacher training and the supervision of teachers to propose that these same observational techniques might be used to good advantage in pre-service and in-service training of teachers. This proposal raises several problems that merit at least brief discussion.

The rationale for using classroom observation systems in teacher training and in supervision of teachers is set forth by Simon and Boyer (1967):

> First, the systems provide a mirror for the teacher to obtain feedback about his own teaching behavior along the dimensions of the particular system used. This feedback provides the teachers with the opportunity to change their own behavior based on data about what they are doing in the classroom. Second, and perhaps more important, many of these systems have been constructed along a theoretical dimension which includes behaviors which are presumed to be helpful in promoting pupil growth if used in the classroom, but which are not ordinarily found in the classrooms of America today. When a teacher uses one of these systems, he gets feedback about the behaviors which he is <u>not</u> using, as well as those which he is. This supplies the chance to learn new behaviors and thus expand the teacher behavior repertoire in ways not ordinarily available to teachers (pp. 19-20).

Few would deny the importance of making available to teachers-in-training and teachers-in-service conceptual tools that enable them to analyze the teaching process and thus to deepen their understanding of the teacher's professional role. If the various systems of analysis developed in the context of research studies are introduced to teachers as intellectual instruments that enable them to view teaching from contrasting perspectives, then teachers may derive great benefit from such

exercises in analysis. Indeed, a recent report by The American Association of Colleges for Teacher Education (1968) proposes that one phase of the training of teachers be devoted to such "analytical study of teaching."

However, a word of caution is in order. The temptation to convert descriptions of teacher behavior into prescriptions for teachers to follow must be assiduously avoided, for few correlates with student learning have been firmly established for the vast array of classroom variables included in existing classroom instruments. Therefore, although observers using the various systems of analysis can provide the teacher with reliable descriptions of many difficult dimensions of classroom events, they obviously cannot prescibe how teachers ought to teach on the basis of these descriptions alone. For decisions regarding what classroom procedures ought to be followed depend both on knowledge of the consequences of various procedures and on judgment regarding the value of these consequences.

References

American Association of Colleges for Teacher Education. Professional Teacher Education. Washington, D.C.: The Association, 1968

Aschner, M. The analysis of verbal interaction in the classroom. In A. Bellack (Ed.), Theory and Research in Teaching. New York: Teachers College Press, Columbia University, 1963. pp. 53-78.

Bellack, A.A., Kliebard, H.M , Hyman, R.T , and Smith, F.L. The Language of the Classroom. New York: Teachers College Press, Columbia University, 1966.

Biddle, B.J. Methods and Concepts in classroom research Review of Educational Research, 1967, 37, 337-357

Biddle, B.J. and Adams, R.S. An Analysis of Classroom Activities. Columbia: Center for Research in Social Behavior, University of Missouri, 1967.

Flanders, N.A. Teacher Influence, Pupil Attitudes, and Achievement. Washington, D C.: U S. Government Printing Office, 1965.

Gage, N.L. Teaching methods, In R. Ebel (Ed.), Encyclopedia of Educational Research, Fourth Edition (in press).

Galloway, C.M. An Exploratory Study of Observational Procedures for Determining Teacher Nonverbal Communication. Unpublished doctoral dissertation, University of Florida, 1962.

Heyns, R.W. and Lippitt, R Systematic observation techniques. In G. Lindzey (Ed.), Handbook of Social Psychology. Cambridge, Mass.: Addison-Wesley Publishing Co., 1954. pp. 370-404.

Hughes, M. Development of the Means for the Assessment of the Quality of Teaching in Elementary Schools. Salt Lake City University of Utah Press, 1950.

Jackson, P.W. Teacher-pupil communication in the elementary classroom: an observational study. Unpublished paper presented at the American Educational Research Association meeting, 1965.

Joyce, B. and Harootunian, B. The Structure of Teaching. Chicago: Science Research Associates, 1967.

Kaplan, A. The Conduct of Inquiry. San Francisco, California: Chandler Publishing Company, 1964.

Kemeny, J.G. and Snell, J.L. Mathematical Models in the Social Sciences. Boston: Ginn & Co., 1962.

Kounin, J.S. An analysis of teachers' managerial techniques. Psychology in the Schools, 1967, 4, 221-227.

Kuhn, T.S. The Structure of Scientific Revolutions. Chicago; University of Chicago Press, 1962.

Medley, D.M. and Mitzel, H.E. A technique for measuring classroom behavior, Journal of Educational Psychology, 1958, 49, 86-92.

Medley, D.M. and Mitzel, H.E. Measuring classroom behavior by systematic observation. In N. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally and Co., 1963, pp. 247-328.

Meux, M.O. Studies of learning in the school setting. Review of Educational Research, 1967, 37, 539-562.

Oliver, D.W. and Shaver, J.P. Teaching Public Issues in the High School. Boston: Houghton Mifflin Company, 1966.

Perkins, H.V. A procedure for assessing the classroom behavior of students and teachers. American Educational Research Journal, 1964, 1, 249-260.

Perkins, H.V. Classroom behavior and underachievement. American Educational Research Journal, 1965, 2, 1-12.

Remmers, H. H. Rating methods in research on teaching. In N. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally & Co., 1963. pp. 329-378.

Ryan, D.G. Characteristics of Teachers. Washington, D.C.: American Council on Education, 1960.

Simon, A. and Boyer, E.G. Mirrors for Behavior, Vol.I Philadelphia, Pa.: Research for Better Schools, Inc., 1967.

Smith, L.M. and Geoffrey, W. The Complexities of an Urban Classroom. New York: Holt, Rinehart and Winston, Inc. 1968.

Smith, B.O. and Meux, M.O. A Study of the Logic of Teaching. Urbana, Ill.: Bureau of Educational Research, University of Illinois, 1962.

Taba, H , Levine, S. and Elzey, F. Thinking in Elementary School Children San Francisco: San Francisco State College, 1964.

Taba, H. Teaching strategies and cognitive functioning in elementary school children. San Francisco: San Francisco State College, 1966.

Wright, E.M. Development of an instrument for studying verbal behaviors in a secondary school mathematics classroom. Journal of Experimental Education, 195?, 28, 103-121.

Withall, J. Development of a technique for the measurement of socio-emotional climate in classrooms. Journal of Experimental Education 1949, 17, 347-361.